

CAPITULO 4

Procesamiento de señales de voz

Continuar

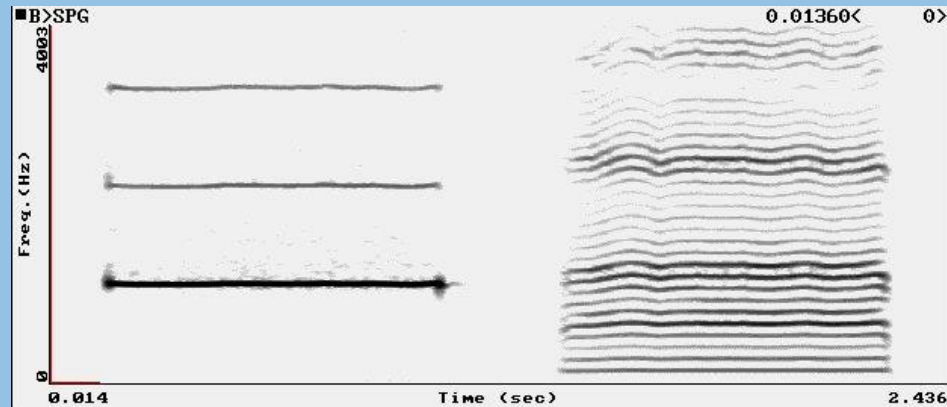
Introducción

Una de las señales más importantes en las telecomunicaciones es la voz. La telefonía tiene como objetivo comunicar a las personas remotamente a través de un medio de comunicación, inicialmente esto se realizó mediante el canal analógico basado en cable conductor de cobre y que ha evolucionado a diferentes medios de comunicación digital, empleando fibra óptica el día de hoy.



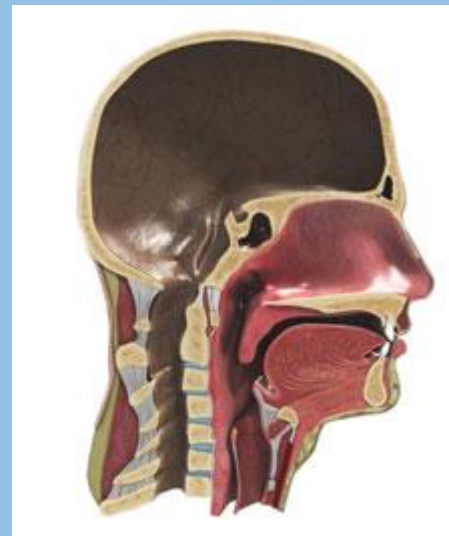
Acústica de la VOZ

Las señales acústicas se forman mediante el desplazamiento de las partículas del aire cuando son excitadas mediante un impulso que genera presión sobre ellas. Las señales acústicas de voz se producen en un rango de frecuencias que varía desde aproximadamente 50 kHz hasta 5kHz.



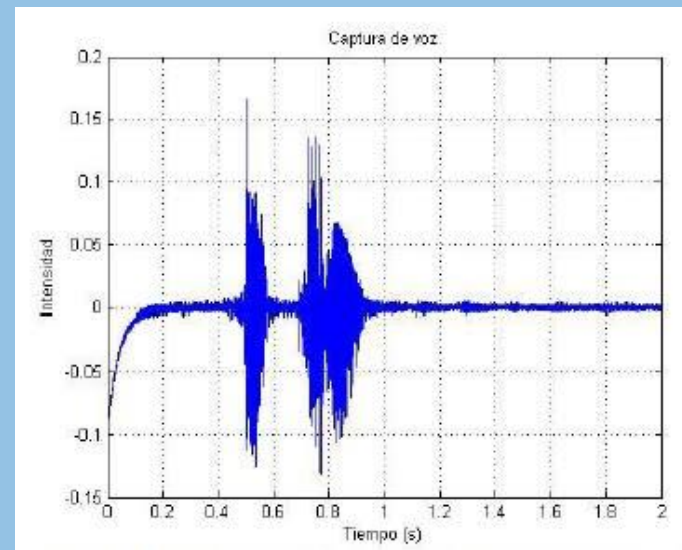
Características de los órganos que integran la VOZ

Los órganos encargados de producir la voz integran el tracto vocal, mismo que puede ser observado como un sistema lineal e invariante en el tiempo al cual se le aplica una señal de impulso a la entrada y a la salida. Una característica importante de la señal de voz es que presenta una frecuencia fundamental y frecuencias armónicas que generalmente son de menor amplitud.



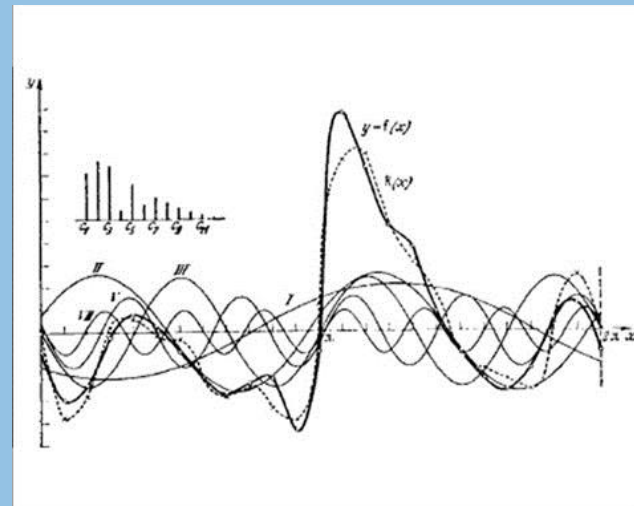
Captura de la voz en MATLAB

El primer paso en el análisis de la voz es la captura de una señal de voz. Mediante el uso de los recursos de la computadora MATLAB permite grabar fragmentos y señales de voz e importar dichas señales al ambiente de trabajo del programa.



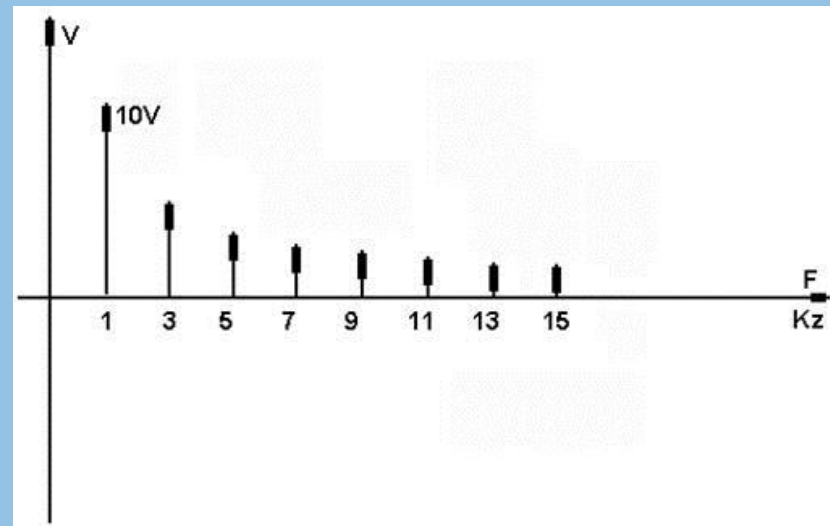
Análisis en el dominio del tiempo

Para el análisis de una señal de voz es necesario dividirla en segmentos de voz para el estudio individual de cada vocal o consonante. Este proceso se realiza mediante la aplicación de las denominadas ventanas que muestran un segmento de la señal original de manera similar a las ventanas empleadas para el diseño de los filtros tipo FIR.



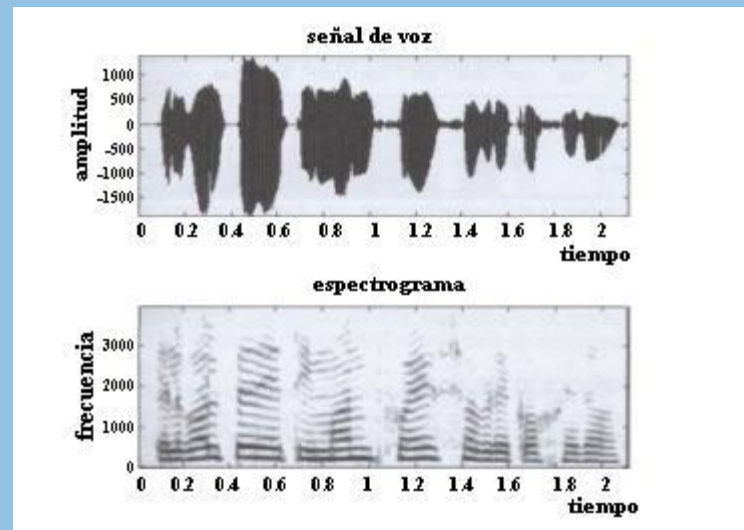
Análisis en el dominio de la frecuencia

La importancia del estudio de la voz en el dominio de la frecuencia radica en la determinación de las frecuencias formante de la voz, las formantes son los picos más importantes de frecuencia de cada sonido y generalmente son 3 los más representativos. Estas forman patrones que nos permiten diferenciar las vocales y reconocer una vocal aún cuando ha sido pronunciada por diferentes hablantes.



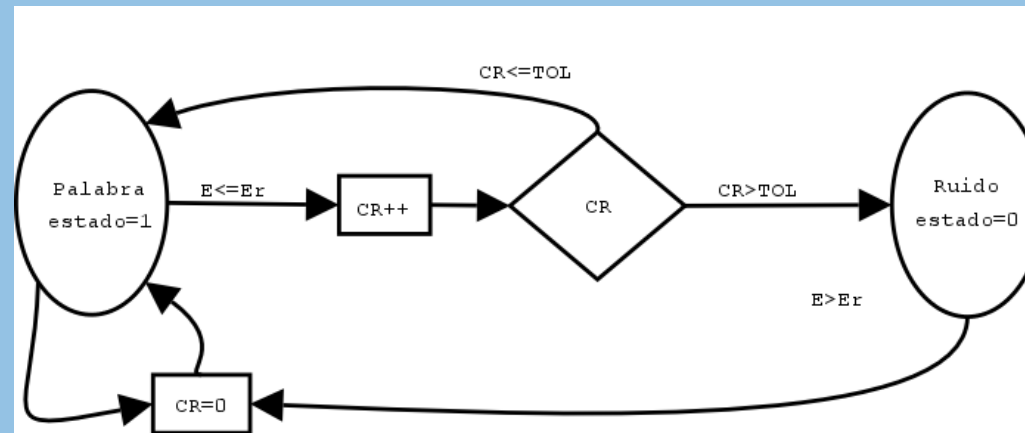
Espectrograma de la señal de VOZ

Un análisis que comúnmente se realiza a las señales de voz, especialmente cuando se trata de frases mucho más largas que una sola vocal es el espectrograma, es decir, una figura en 2-D donde se muestra el contenido de frecuencia de una señal y como varía en el tiempo.



Predicción lineal de la VOZ

Se analiza la técnica para análisis de la voz denominada predicción lineal. El problema de predicción lineal tiene como finalidad la estimación de los parámetros de un modelo que represente la señal de voz que se está escuchando.



Teoría de estimación

Mediante la llamada teoría de estimación existen dos aproximaciones para la solución del problema de determinar coeficiente, una es la teoría clásica que asume que θ es una variable determinística y por otro lado se encuentra la teoría Bayesiana en donde considera que esa misma variable es aleatoria. Propiedades que debe de tener el estimador:

- La media del estimador es el valor real del parámetro.
- La variancia del estimador debe ser mínima.

Parámetro poblacional	Estimador	Estimación
Media μ	$\hat{\mu} = \bar{X} = \frac{\sum_{i=1}^n X_i}{n}$	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$
Varianza σ^2	$\hat{\sigma}^2 = S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$	$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$
Proporción p	$\hat{p} = \frac{X}{n} = \frac{\text{número éxitos}}{\text{número pruebas}}$	$\hat{p} = \frac{x}{n}$